Poke Data

Process Book

Kaelin Hoang and Sunny Siu

Table of Contents

Overview and Motivation

Data visualization is commonly used for communicating more serious things like scientific findings, historical trends, or current events. We felt as though there's an influx of negative current events so we wanted to do something that was more lighthearted. Therefore we decided to use a Pokemon dataset in order to create interactive, yet still educational, visualizations with Pokemon data. The show has a handheld computer called a Pokedex, which when pointed at a Pokemon, explains everything about it. This always felt so magical to us, so we wanted to bring that item to life, minus the picture taking and talking aspects.

Related Work

The original pokemon colors were too saturated and bright for the aesthetic we were going for. For the overall page we went with the colors of a pokeball(red, white, black) since this was pokemon afterall. As for the type colors, we found another pokemon dashboard with more muted colors for each type and used that. The colors can be seen below. (http://jkunst.com/flexdashboard-highcharter-examples/pokemon/).
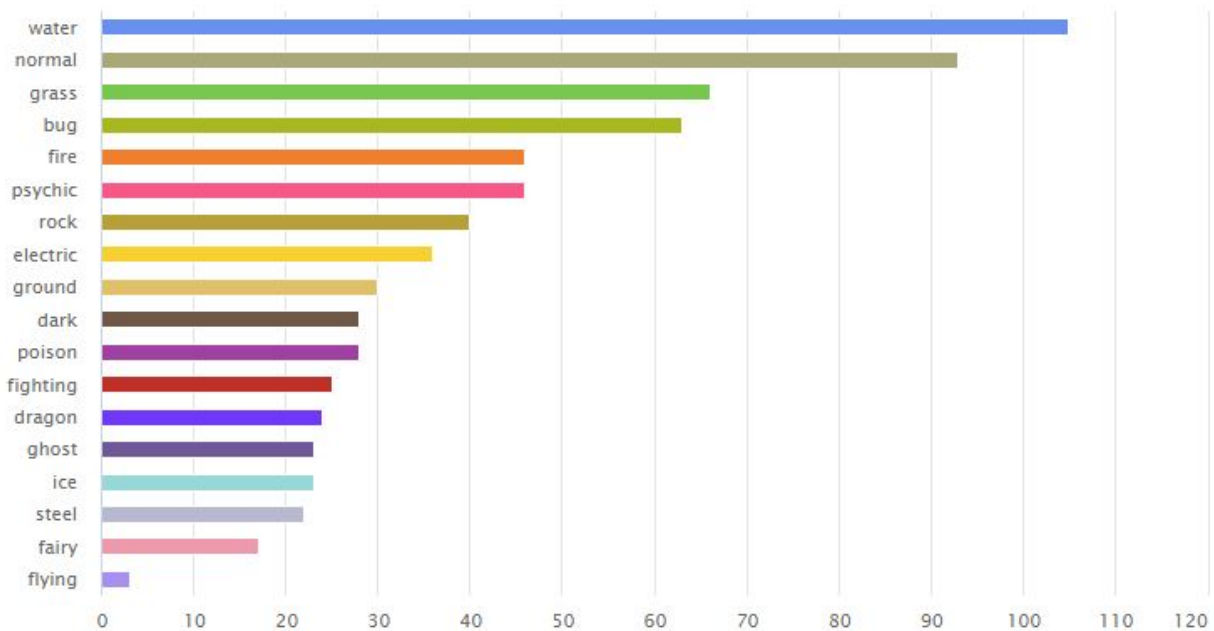
Questions

This project aims to create an interactive and educational data visualization for all 8 generations of Pokemon. We hope that it will help give a better understanding of pokemon statistical distributions in terms of types, generation, and stats. There are so many pokemon and type-relations that we want someone unfamiliar with the topic to be able to gain an understanding of the pokemon ecosystem. Also for people who already know about the pokemon world, this tool can be used to find out more in depth relations, like how strong certain pokemon are in relation to the rest.

Data

The Complete Pokemon Dataset (https://www.kaggle.com/rounakbanik/pokemon) from Kaggle is accessible through a downloaded csv file. In order to get the sprites for each pokemon, we will be using the PokeAPI's github (https://github.com/PokeAPI/sprites/tree/master/sprites/pokemon) which has a folder of all Pokemon images based on that Pokemon's Pokedex ID.

Data Update

One of the features we tweaked was how we should show the evolutions of pokemon. We realized the original Kaggle dataset did not have this information. We were able to find a csv with the evolution data from this public Github repository (https://github.com/jenniNelson/dataviscourse-pr-christennelsenson/tree/master/WhosThatPokemon/data/pokemon_data) and used that.
After loading in the evolutions we noticed that many of the arrays and values were the wrong types, so we wrote a python script to fill the null values and convert the fields to the correct data types. We also converted the csv files to json as that sped up the load time a little bit.

An extra feature we decided to add was an arc plot to show relationships between pokemon types. This new feature required data about the relationships between types, which we found from the PokeAPI here (https://pokeapi.co/docs/v2#types). A new python script, types_collector.py, to call the API and format this data into a json file, types.json, that lists what each type is strong against, weak against, and has no effect against. We then used d3.chords to convert this into edges and nodes.

Exploratory Data Analysis

The data for this project was found in csv files which is easy to analyze when opening in Excel. By scrolling through the tabular data, we were able to deduce what kind of visualizations we could create with the data given. Additionally, the data is built on our previous knowledge so we could make assumptions on what we wanted a lot of the information to look like when displayed.

Design Evolution

Initially, we came up with 3 design sketches that each had some unique strengths. We then combined the best ideas from the 3 designs into a 4th final design. We used this final design as a base for starting our project and adjusted it from there.

Upon selection of a Pokemon, we want to be able to show a breakdown of their stats (defense, attack, etc), type(s), height, weight, and any other classifying information. In order to display pokemon stats, we want to utilize a radar chart to display special defense, special attack, defense, attack, and speed.

Prototype 1
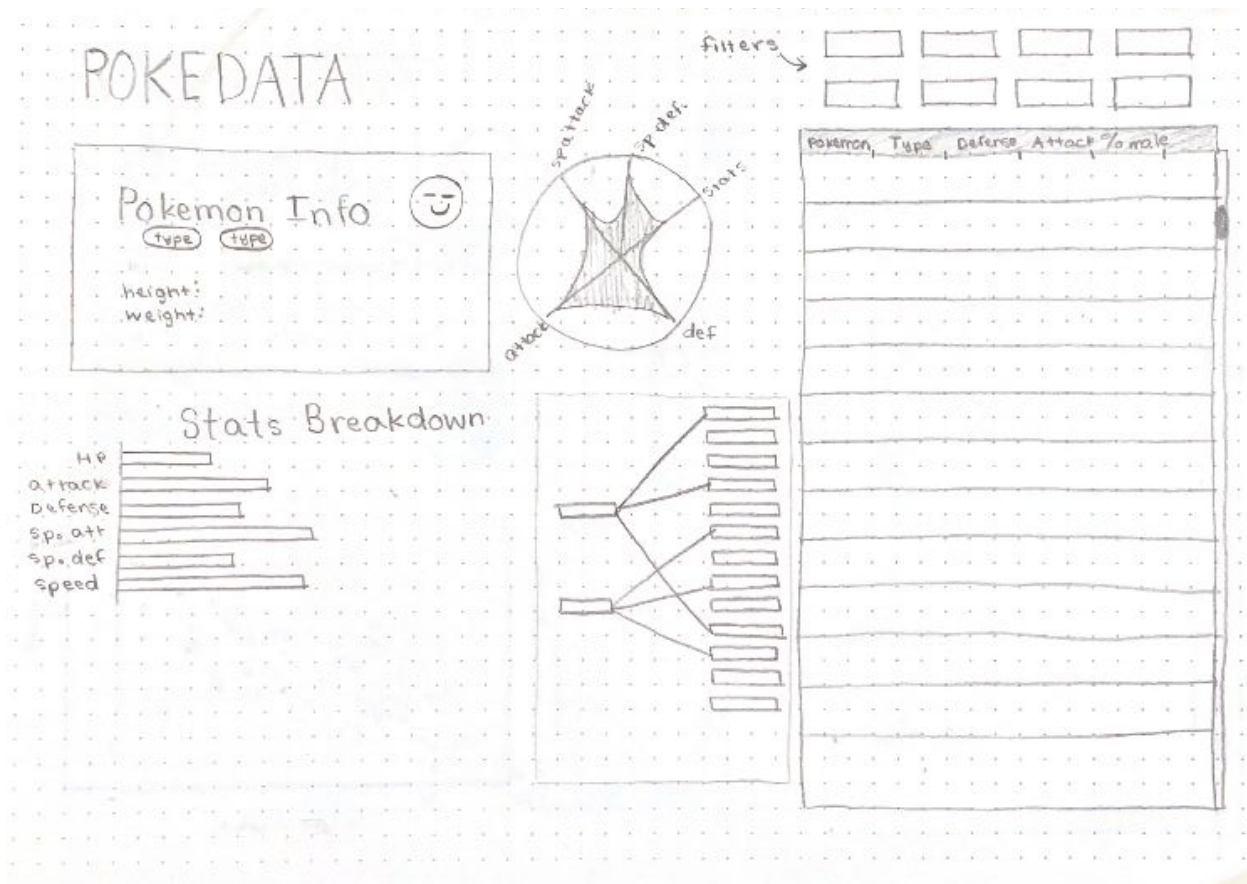
The first prototype design features a table view on the right hand side of the screen which will have every pokemon listed along with all of its stats. We also want a way to filter through all the pokemon in the table, hence the filtering buttons at the top of the table which could filter on type. Then on selection of a pokemon within the table, it would pull up a breakdown of that pokemon's statistics by use of an info card and a few different graphs. The info card contains the name, picture, type(s), height and weight. We wanted to show a visual indicator for specific stats (defense, attack, special defense, special attack, and speed) in a traditional radar chart. This way allows for a quicker understanding of how each stat relates to one another. Then we break down the rest of the statistics a little further in a stats breakdown chart which is a simple horizontal bar chart. To the right of the stats breakdown is information about the selected pokemon's type and what types the pokemon is weak and strong against. We envisioned this would be a visual indicator similar to word-term matching on tests where there are lines to types that are more or less effective given the type of the selected pokemon.
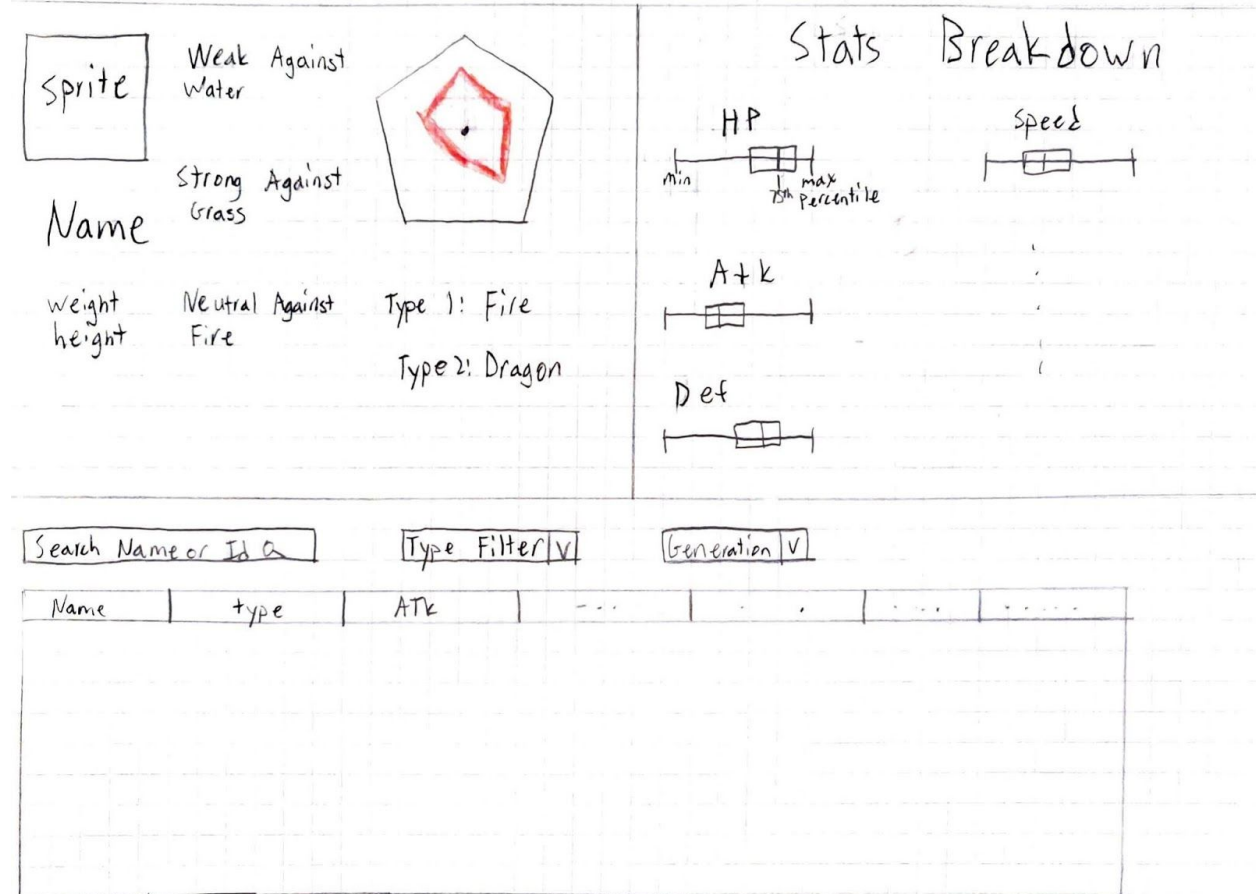
Prototype 2
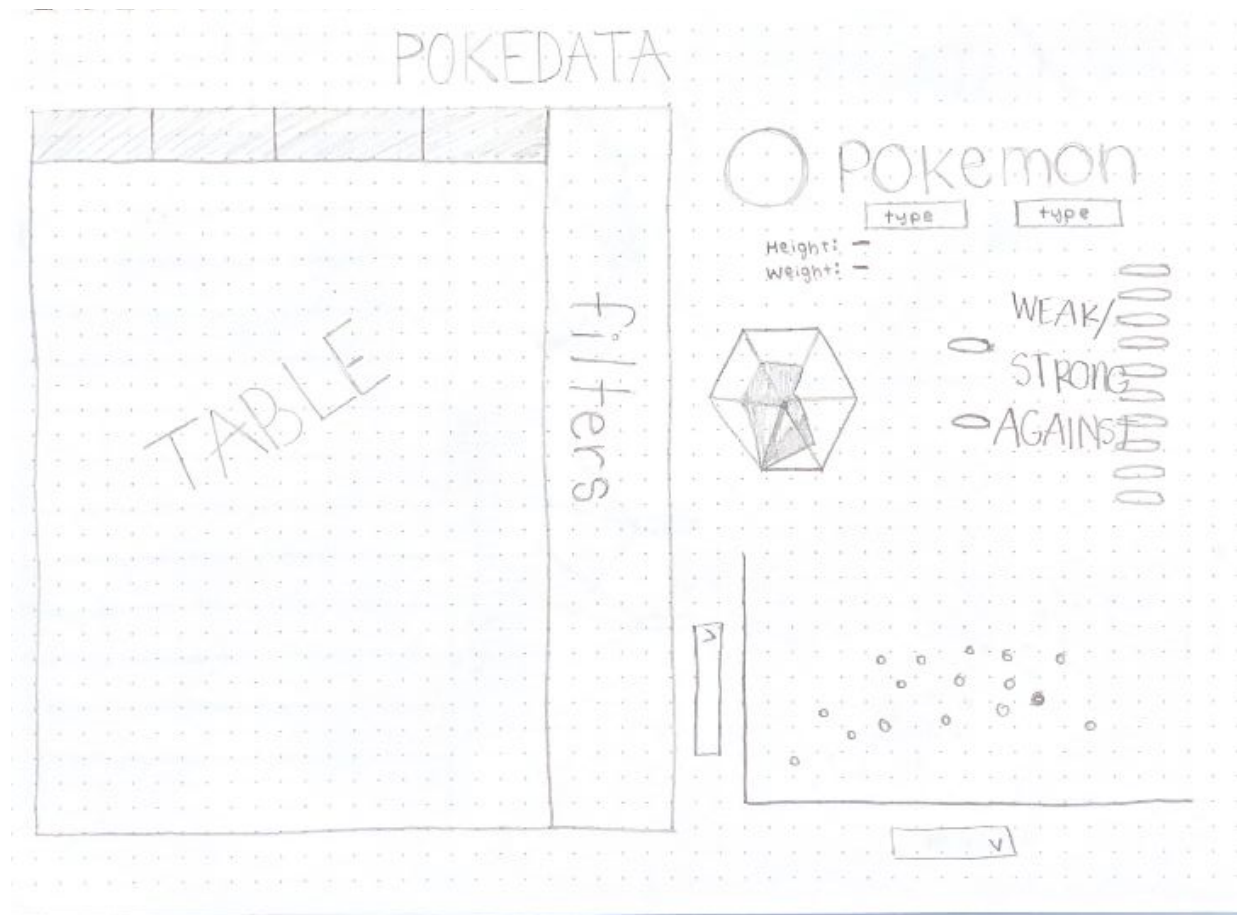
The second design is similar to the first, but the major differences are what is the most highlighted aspect of what is seen first. The table is a big focal point of this design because we have a larger table on the left which has a filter bar on its right hand side. Interactions within the web page would work similarly as selection of a pokemon in the table would pull up it's statistics breakdown on the top right side. The top right side still has the info card, radar chart for stat comparison, and then the weak against and strong against type information for that selected pokemon. Here we wanted to add in a distribution overview of all pokemon in the table on the bottom right of the webpage. The scatter plot has customizable x and y axis statistics which would display every pokemon that is in the table. The plot would filter as the table filters as well. We thought this would be a useful change to not only have an individual pokemon's statistical information but to also compare that pokemon with all the rest of the pokemon as well.
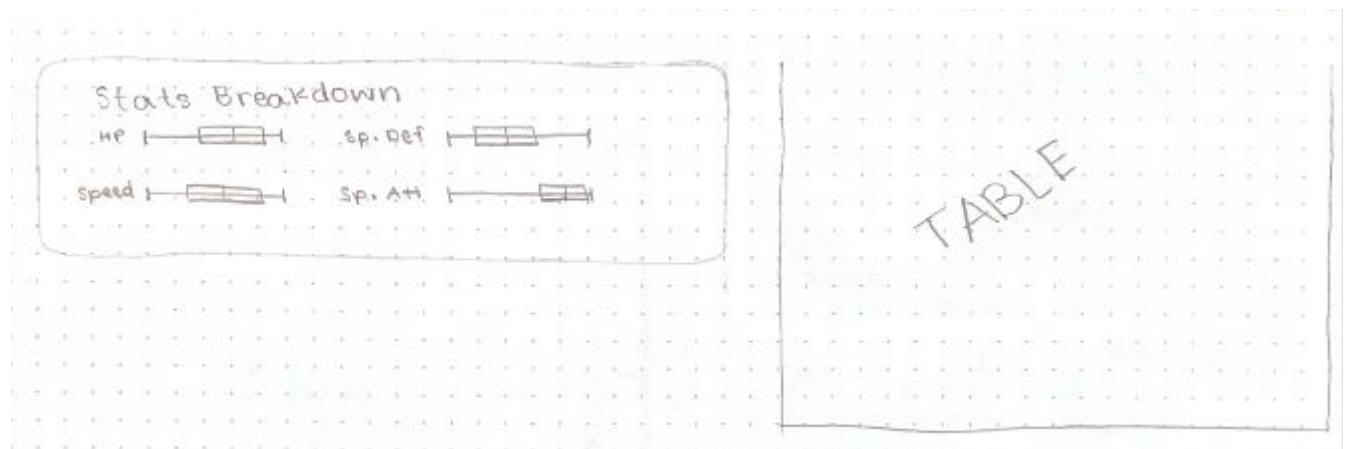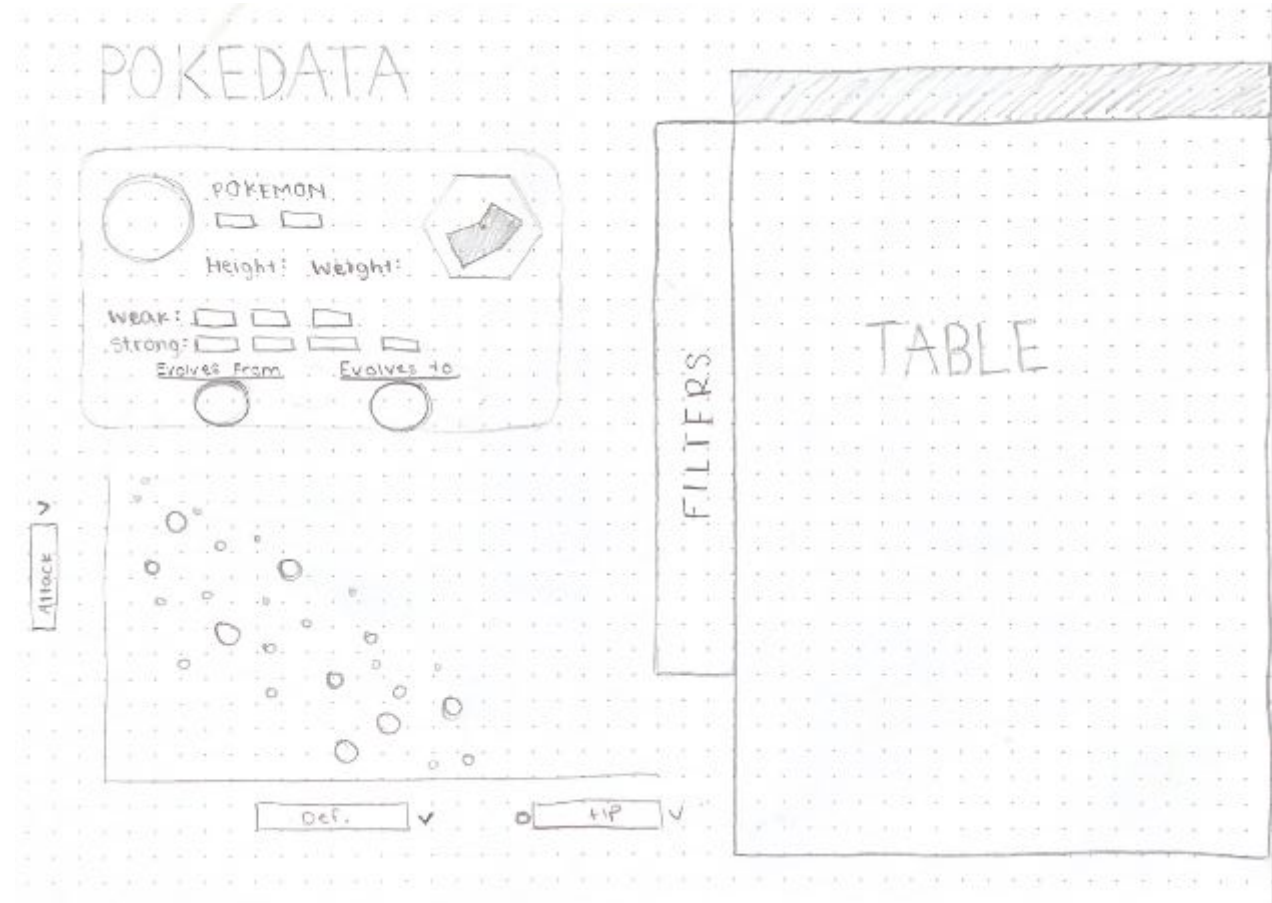
Prototype 3

Pokedata
Authors

Sprite	Weak Against Water			**Stats Breakdown**
Name	Strong Against Grass			HP
weight height	Neutral Against Fire	Type 1: Fire		
		Type 2: Dragon		Speed
				Atk
				Def

HP — min / max / 75th Percentile

Speed

Atk

Def

Search Name or Id

Type Filter ∨

Generation ∨

Name	type	ATk				

This design splits our project into 3 sections: Infocard, stats breakdown, and table. The table gives an overview of all the data, since the data is tabular in nature this felt like the natural thing to do. This allows for easy filtering and ordering of the data as there are too many rows to digest at once. The top half is a zoomed in view for the currently selected pokemon. For easy comparing of the stats we drew a radar chart that will transition on selection. We felt this was the easiest way to compare multiple stats compactly (attack, defense, speed, hp, special attack, special defense) as it encodes each value with their distance from the center. As for the stats breakdown, we wanted to show how the selection compared to the dataset overall. Since there are 6 stats, we wanted a compact way to display what percentile a pokemon's stats fell into.

Final Design

POKEDATA

POKEMON

Height: Weight:

Weak:
Strong:
Evolves From.   Evolves to

Attack

Def.   HP

FILTERS

TABLE

Stats Breakdown
HP   sp. Def
Speed   Sp. Att

TABLE

We designed each prototype with a specific strength in mind, but realized we do not have to confine our visualization to a single page because you can scroll down infinitely in a webpage. So we combined all the elements in each prototype into this once. The most important part of a pokedex is finding the pokemon you want and seeing its core information, so we put the table and infocard at the top. The table lets a user quickly sort through the tabular data by any attribute (name, type, stats) as well get an overview of all the data. The infocard will include the stats in a radar chart, which we discussed compactly encodes stat values with distance. We then added a scatter plot below the infocard so that users can compare any 3 stats on the axes. Since position and area are very powerful channels we believe this is the best way to compare a few specific attributes. Unseen here, but the color will also encode the pokemon's type. Clicking a point will bring it up on the infocard as well as in the stats breakdown, which lets the user select a pokemon with specific criteria in mind. The stats breakdown will use boxplots to show how a pokemon's stats compare to the rest. We had a lot of stats so we wanted something compact and easy to read quickly.

Implementation
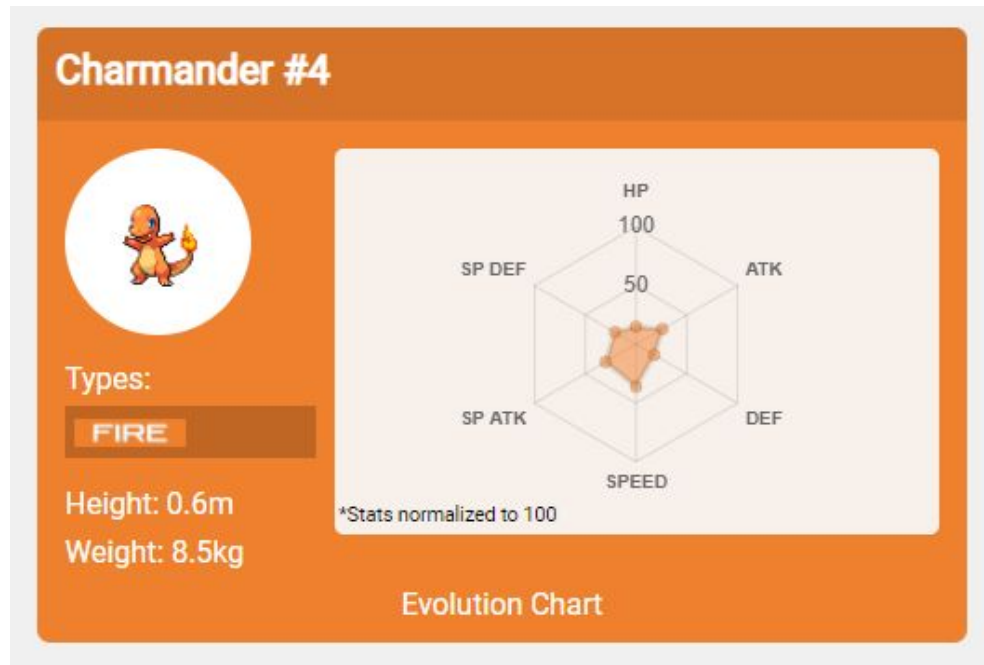
Overall Layout
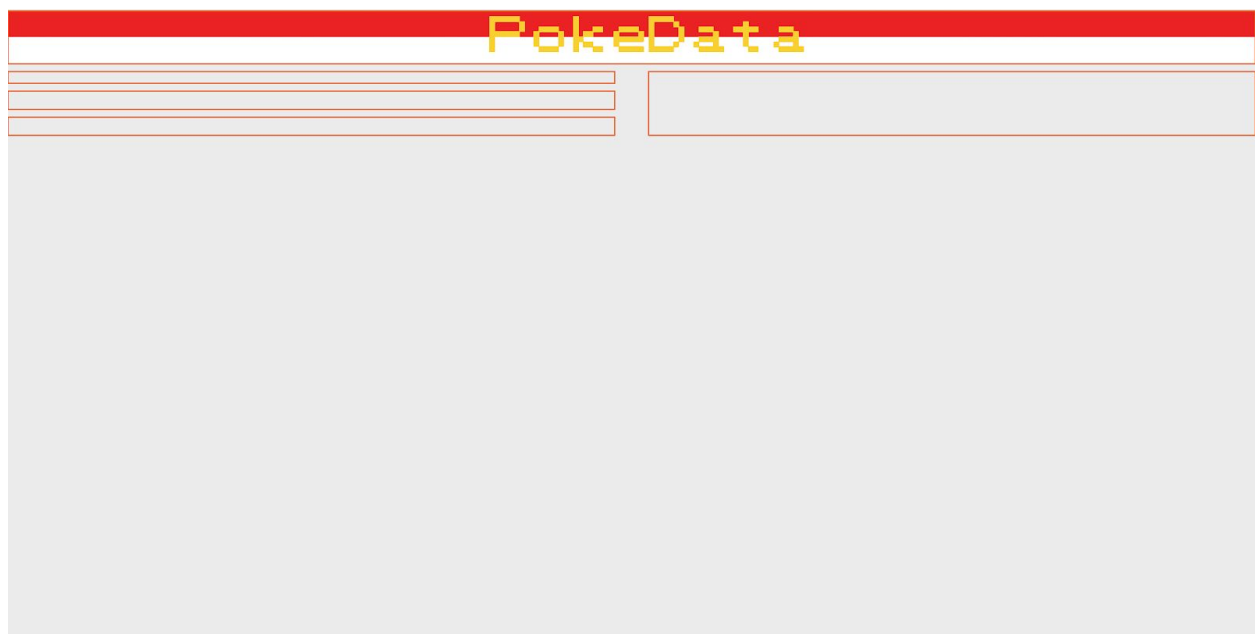
Our base layout includes a title banner, complete with boxes for where we plan to put our components. The left side will have the infocard, scatterplot, and stats breakdown, while the

right will have the table. We felt this was a good starting point as it separates all the components of our project and allows us to split the work.

Infocard

The infocard is meant to give fine control to a user by showing the attributes of a selected pokemon. The sprites are unique and the core of the franchise, but there are over 800 of them so it is important to allow users to map the image to the name. We also wanted a way to convey the type to the user, so we have a double encoding for it (the color of the infocard and the actual logo in the types section). A small dark background was applied to make the type logo discernable from the background itself. As for the stats window, we wanted to give a quick overview at a pokemon's best and worst stats. Because HP has a much bigger domain than the other stats, they were all normalized to 100 to keep every stat visible. The radar transitions to the new size when a new pokemon is selected, as opposed to no transition, which allows users to visually compare the stats.

Charmander #4

Types:

FIRE

Height: 0.6 m
Weight: 8.5 kg

*Stats normalized to 100

Stat chart: HP, ATK, DEF, SPEED, SP ATK, SP DEF (values 50–100)

Evolution Chart

Evolutions are another key component to pokemon and creates a relationship between them. Pokemon typically start in a small form and as they get stronger (or a special item is given to them) they will evolve into another pokemon. Since this is a hierarchical process (pokemon always evolve from one to another and cannot go back) we thought a tree would show this best. Each node encodes the pokemon with its sprite, and clicking any of the sprites will select that pokemon and change the infocard accordingly.

Table

#	Name	Type 1	Type 2	HP	Attack	Defense	Special Attack	Special Defense	Speed
1	Bulbasaur	grass	poison	45	49	49	65	65	45
2	Ivysaur	grass	poison	60	62	63	80	80	60
3	Venusaur	grass	poison	80	100	123	122	120	80
4	Charmander	fire		39	52	43	60	50	65
5	Charmeleon	fire		58	64	58	80	65	80
6	Charizard	fire	flying	78	104	78	159	115	100
7	Squirtle	water		44	48	65	50	64	43

The first version of the complete pokemon table was a very simple, scrollable table. The table mimicked the csv file with no particular changes to the display of the data, other than the ability to sort the data by clicking on the headers. By default, the data was already sorted by Pokedex number. At this stage, selection of a row will also update the infocard with that pokemon's picture and stats.

#	Name	Type 1	Type 2	HP	Attack	Defense	Special Attack	Special Defense	Speed
1	Bulbasaur	grass	poison	▪	▪	▪	▪	▪	▪
2	Ivysaur	grass	poison	▪	▪	▪	▪	▪	▪
3	Venusaur	grass	poison	▪	▪	▪	▪	▪	▪
4	Charmander	fire		▪	▪	▪	▪	▪	▪
5	Charmeleon	fire		▪	▪	▪	▪	▪	▪
6	Charizard	fire	flying	▪	▪	▪	▪	▪	▪
7	Squirtle	water		▪	▪	▪	▪	▪	▪

The next iteration of the table was to visually encode how the individual stats were displayed. We thought that plain numbers were not as intuitive or comparable so we decided to add in the bar visualization of the data. Additionally, each bar is scaled with the minimum and maximum of each individual stat to better display the various ranges. A tooltip will display the actual value of the stat when hovered over as well to still convey the same information as the previous version.

#	Name	Type 1	Type 2	HP	Attack	Defense	Special Attack	Special Defense	Speed
1	Bulbasaur	GRASS	POISON						
2	Ivysaur	GRASS	POISON						
3	Venusaur	GRASS	POISON						
4	Charmander	FIRE	--						
5	Charmeleon	FIRE	--						
6	Charizard	FIRE	FLYING						
7	Squirtle	WATER	--						

For more consistency throughout the entire webpage, we thought it would be a good idea to also encode the types using the type badges used in the infocard. For this version of this table, filters were added in order to make looking for a specific pokemon a little easier or to compare pokemon of a similar type/stat range. The search bar searches on name and pokedex number while the 'filters' button expands to show type filters and stat sliders.

There are two distinct sections in the filters' expanded view: type filters and range sliders for each stat. Clicking a type filter will filter the table on whether a pokemon has that type for either their first or second type. A visual indication of the filter selection will be represented as

that type button with an 'X' next to it to provide a way to clear the filter. The range sliders for the individual stats will filter based on the numbers selected by each slider.

#	Name	Type 1	Type 2	Attack	Defense	Speed	HP	Special Attack	Special Defense
1	Bulbasaur	GRASS	POISON						
2	Ivysaur	GRASS	POISON						
3	Venusaur	GRASS	POISON						
4	Charmander	FIRE	--						

Color encoding for stats was found to be more important than showing the type color again for that pokemon's stats. Each stat was given a distinct color which helps distinguish between each stat. Additionally, these colors are used differently than how types are used since type colors are specific for the selected pokemon.

Filters ▲

☑ **Legendaries**

WATER NORMAL GRASS
BUG FIRE PSYCHIC
ROCK ELECTRIC GROUND
DARK POISON FIGHTING
DRAGON GHOST ICE
STEEL FAIRY FLYING

Attack: 5 - 185

Defense: 5 - 230

Speed: 5 - 180

HP: 1 - 255

Special Attack: 10 - 194

Special Defense: 20 - 230

Eventually, the ability to filter on whether a pokemon is legendary was added in as one of the final additions to the table. For consistency throughout the website, all legendary pokemon names are marked with a star.
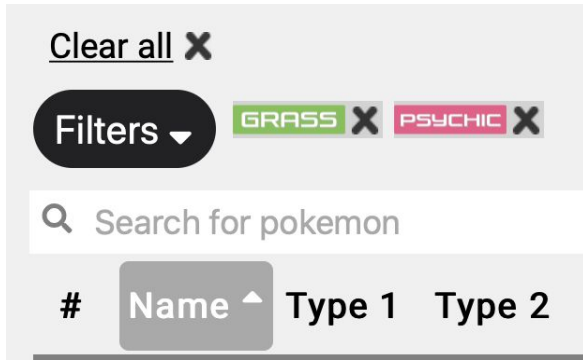
The final step was to make sure that, if the user wanted to clear any filters, that they could do so with the click of a button. By clicking on this button above the table and filters, all filters will clear, stat filter sliders reset to their original state, data will unsort, and the table will be in its original form.

Scatterplot
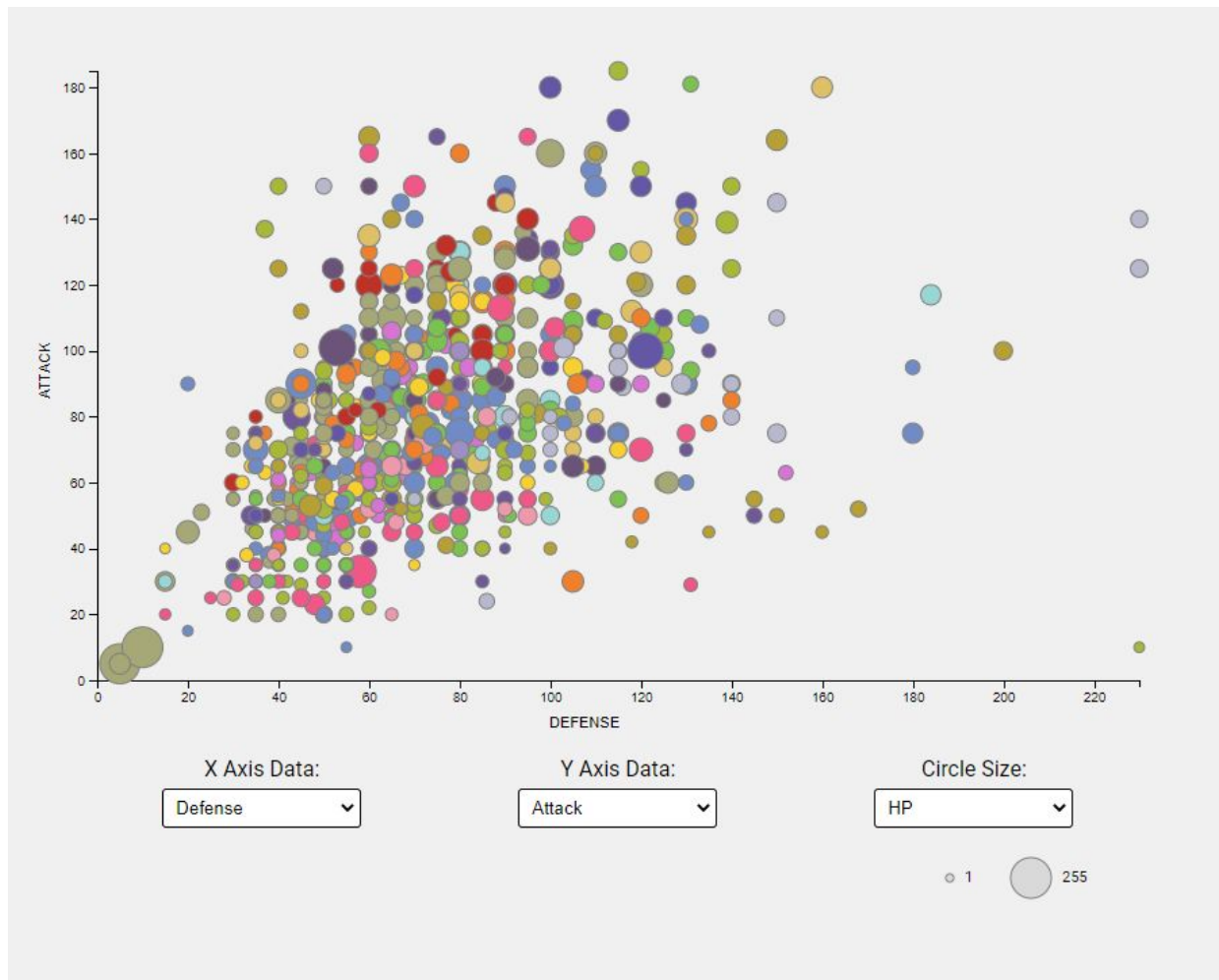
A core mechanic of the pokemon games is the battle system. For this reason, we wanted to compare pokemon by their stats. As seen with the radar, it can be hard to look at all 5 at once, so we allow users to intuitively compare 3 up to stats. Users can select which stats they want on each axis, as well as what they want the circle size to represent. There is also a legend to indicate what values the circle sizes represent as that cannot be seen in the scatterplot. This plot was effective in the Gapminder assignment so it was reused here. Color encodes the type of each pokemon, which is a recurring color encoding that will be seen across all the views.

Rayquaza #384

Types:
DRAGON FLYING

Height: 7 m
Weight: 206.5 kg

*Stats normalized to 100

Evolution Chart
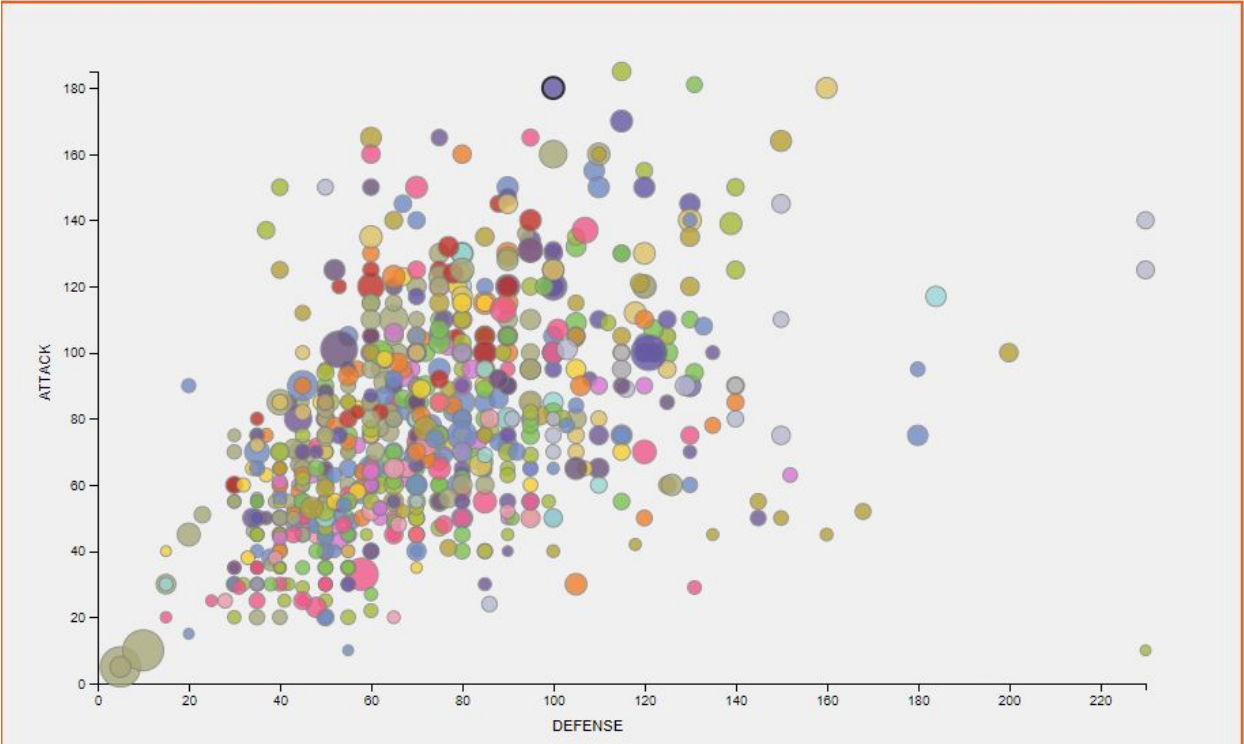
Being able to compare stats alone is useless if we do not let users look at the pokemon each dot represents. So if a user hovers over a dot the tooltip will show the pokemon name and pokedex id (which is a unique id that each pokemon has). They can click the dot to select it and the infocard will update accordingly. This gives the user a way to filter pokemon by certain stats before selecting them. Also worth noting, the opacity of the circle was adjusted in this iteration so that overlapping circles could be seen.
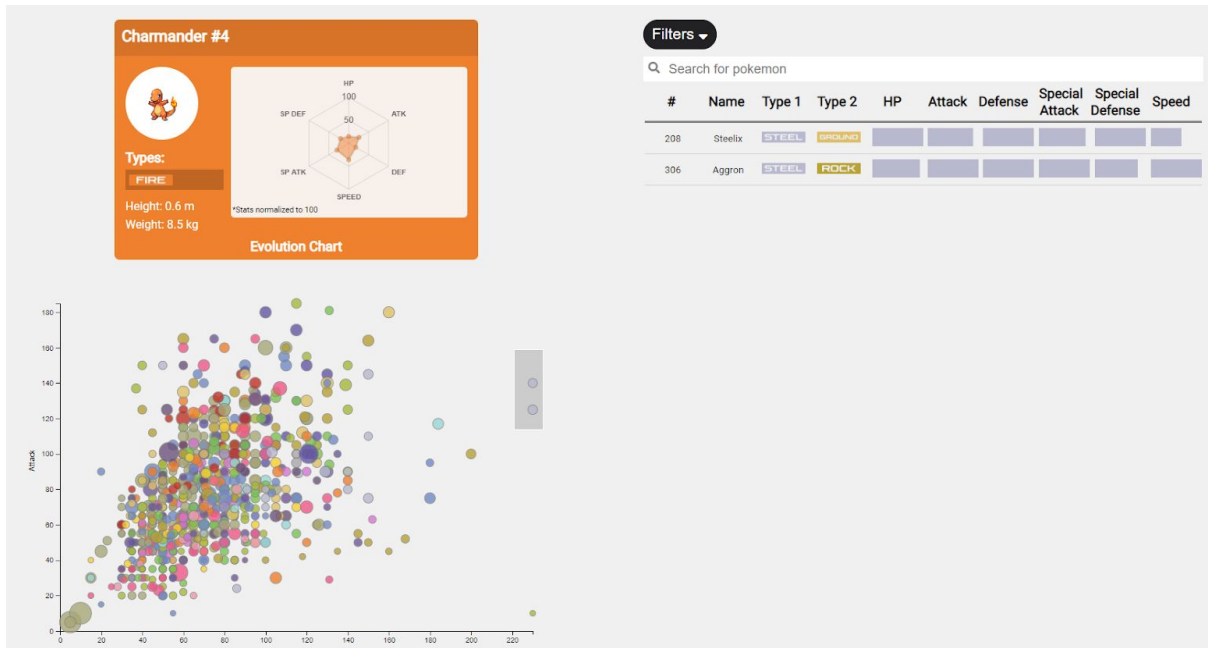
A brush was added to allow users to narrow down specific pokemon they want to look at. Brushed circles will adjust the table accordingly, which gives the user another form of filtering. This was necessary because the infocard can only show 1 pokemon at a time, but the table allows for much more. And because of the bar encodings in the table, it also allows users to compare in the table as well as the scatterplot.

Navbar

A filter was added to our navbar to pick which generation of Pokemon to use for the page. We thought of putting this in the table filters, but as this affected the entire page, it seemed fitting to put it in the banner.
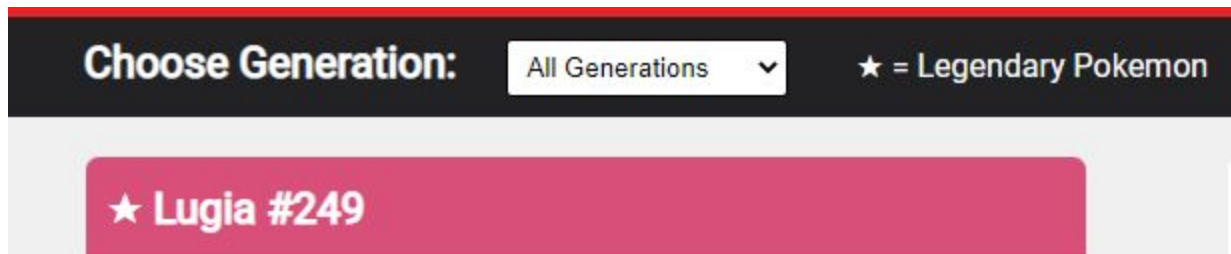

In this iteration we added a universal legend to indicate legendary Pokemon. This star can be seen in the infocard, table, and scatterplot. It is instantly recognizable regardless of shape and color. This helped mark legendaries as they are special in the pokemon world and a common thing to search for.


For the final navbar we added links to our Process Book and Demo Video. The star legend was moved to the right side as it was not interactable and felt awkward being grouped with everything else.
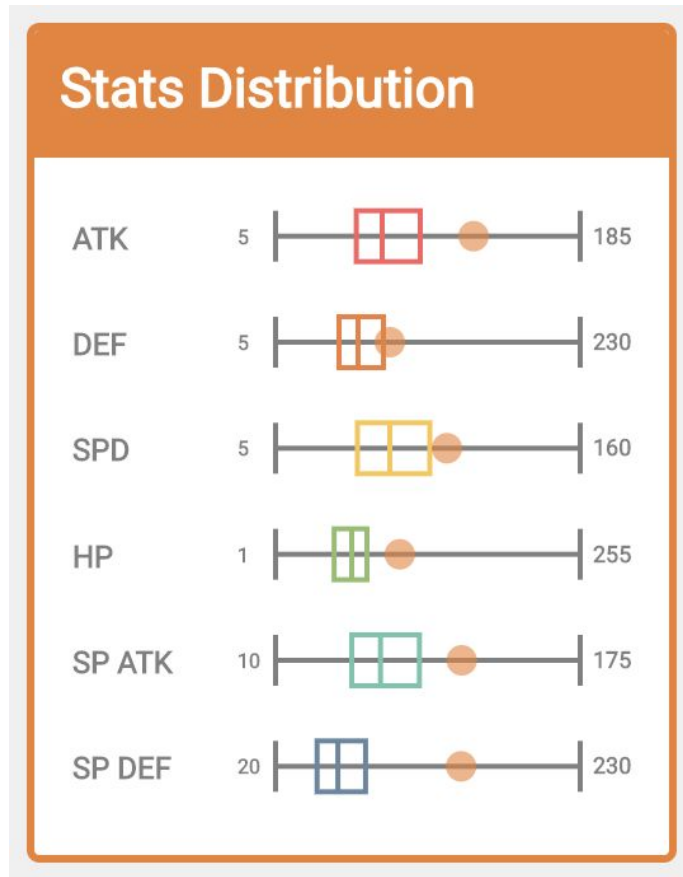
Box Plots
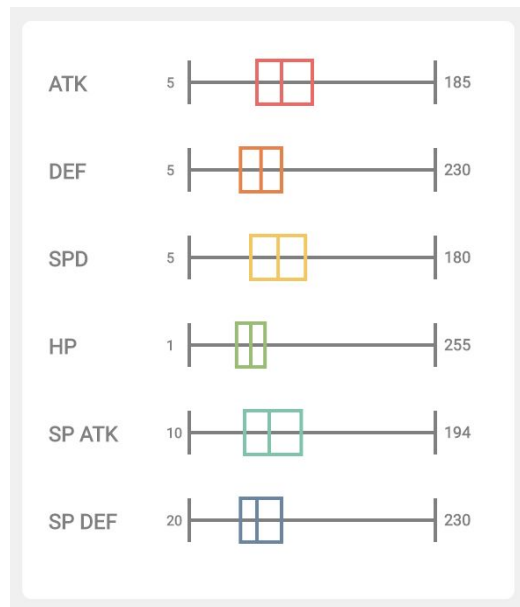
The first version of the box plots displayed the total distribution of all pokemon. These did not get filtered when the table filtered and did not show the current selected pokemon. It was mainly a version showing the general layout of how the box plots should look. Each stat was created to be the same length but since each stat has a distinct range, each plot used a different scaling scheme. The min and max values are indicated by labels at the ends of the plot and there are plans to add tooltips for the middle quartile range values.
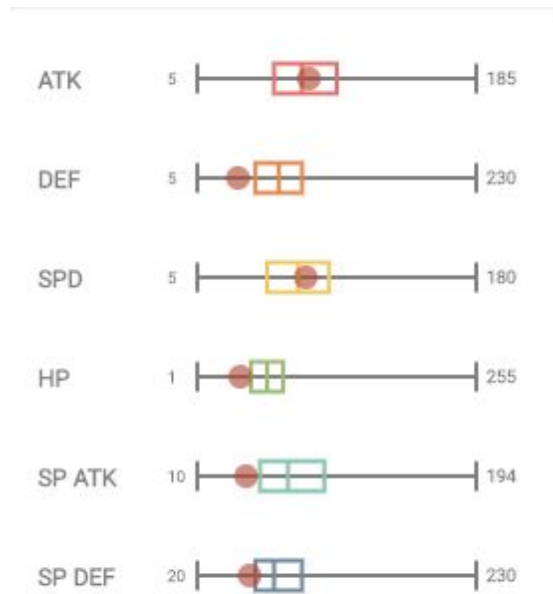
The next version of the box plots spaced the plots out more vertically and added in the current selected pokemon. Although not shown here, upon filtering will change the distributions of the box plots which will update the current selected pokemon's indicators on the plots. The current styling was just a white background on the light gray of the website. It's styling didn't match the infocard but was a clean implementation that would need some more iterations later on.

This image displays the final version of the box plots. With an added title and border (which changes color based on the selected pokemon), there's a better visual distinction between the background of the website that still ties in the visualization nicely with the infocard that it sits next to. With this version of the box plot, it filters and updates the distributions as well as animates the transition of pokemon selections from either the scatterplot or the table. There's a tooltip on hovering over the circle which will show the current value of the selected pokemon for that given statistic.

Type Chart
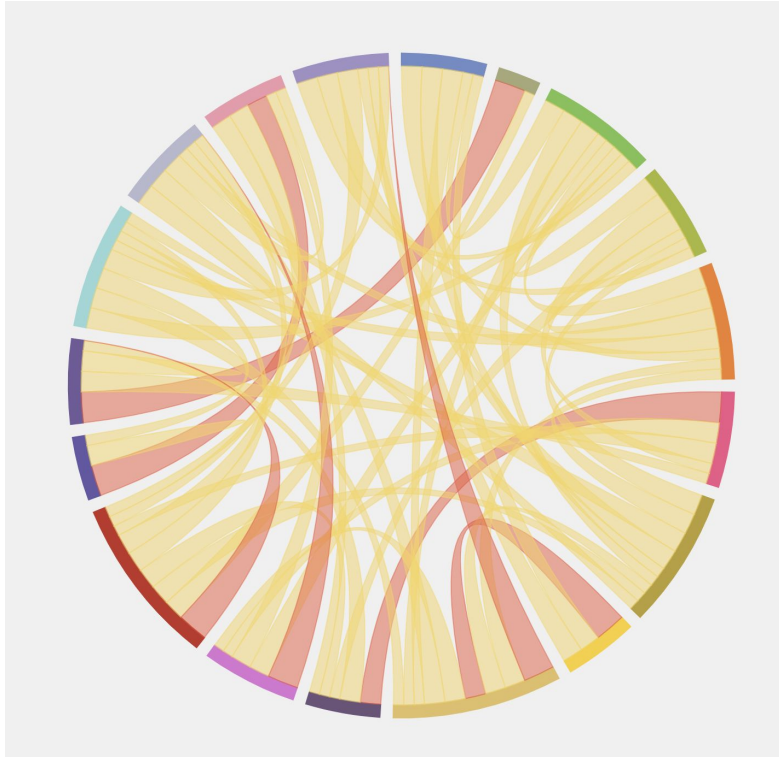
We had some extra time after all the features, so we decided to implement our bonus feature. The idea was to have something flashy to show off relationships between types. Types in Pokemon are like a complicated game of rocker-paper-scissors. For instance, fire beats grass, grass beats water, water beats fire. It gets much more complicated though, as every type has another type it is strong, weak, and has no effect against. Our first thought was to use a chord chart, but while implementing it we realized that chord charts do not work the way we want. First, not all type relationships are two way (ex. ground attacks have no effect against flying, but flying is not strong against ground) and there is no proportions between the two types. We also thought about how this chart was too complicated for a part of our target audience, children.
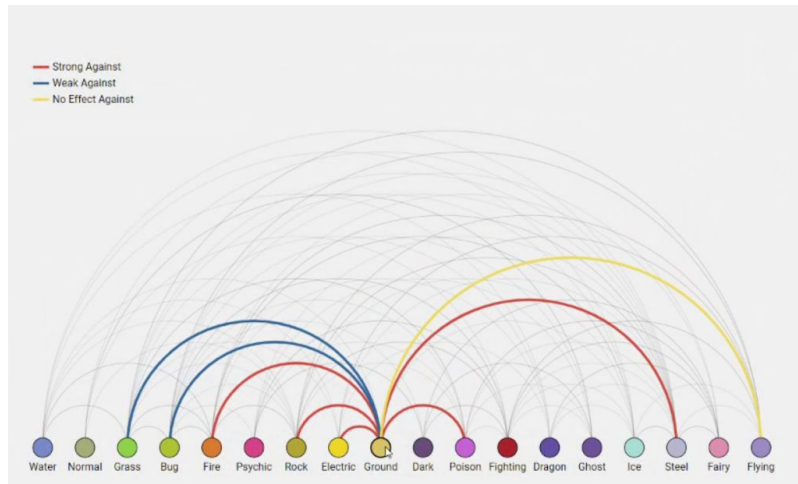
We found an arc plot was simpler and fit our needs. It was flashy, correctly showed relationships without proportion, and did not require two way relationships. When you hover the mouse over a type, it highlights arcs that come out of it and color them according to their relationship attribute. The color code is blue for weak against, red for strong against, and yellow for no effect against. This works for 1-way relationships as well. When you hover over ground, it has a yellow line going to flying, but if you hover over flying it has no line going to ground. The clutter of the lines works to our advantage in this case, because though there is a line going from flying to ground, it is not easily visible and thus softly hidden.

Finally, to add a storytelling aspect, we added a little intro to the type chart to explain how it works and what it shows. An expert in Pokemon could easily skip this and still understand what is going on, but a newbie could read this for a better explanation. We know this works as we sent out this site to our friends and it had the intended effect.

Evaluation

With our visualizations, we were able to see that the ability to compare pokemon was best used with many interactions. Most visualizations for pokemon (other than the type chart) interacted in a way, to show many different ways a pokemon stacks up against others. Filtering on the table allows distributions to be recalculated for the box plots and for circles to filter out on the scatterplot. So with the interactions we provided, the exploration for each pokemon and how it compares to other pokemon is endless.

A constantly difficult problem we faced during the project was picking colors. This was difficult, because the Pokemon types use 18 colors and we use those colors to visually encode many things, but did not want unrelated things to be mistaken for a type encoding. The solutions we came up with were to create clear separations in each place color is used. We also used white/gray/black for neutral things such as dropdowns, buttons, and miscellaneous interactable items.

Throughout our implementation, we have faced design decisions that dictated how the behavior of our final project. Originally, the scatterplot's axes redrew upon filtering the table, but then it was decided that the axes should always be scaled according to all possible pokemon stats to see how filtering affects the pokemon in question. Additionally, stats were encoded with their own colors for distinction in the table bar indicators as well as the box plots. The colors chosen were used in different ways from the type colors in hopes of not having confusion amongst the multiple colorways that are used in the website. Box plots were scaled on each individual distribution for each statistic. The box plot distributions changed upon filtering the table, unlike the scatterplot. The distinction here is now the user can view filtered distributions via the box plots and unfiltered distributions using the scatterplot just below it. Additionally, there's no need to compare the distributions amongst statistics, users care more about comparing the selected pokemon against the rest of the pokemon. In the final prototype design of our website, we thought that we could list out the types of the selected pokemon that it was weak and strong against, However, type weaknesses and strengths don't change by individual pokemon; they are consistent among all pokemon that have that certain type. Therefore we decided that a type chart would be more usable and better to generalize the strengths and weaknesses of each type.